DOCKET NO: 10671-0001-2 PCT

09/462562
416 Rec'd PCT/PTO 10 JAN 2000

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Zigmantas L. BUDRIKIS, et al.

SERIAL NO.: NEW U.S. PCT APPLICATION

FILED: HEREWITH

INTERNATIONAL APPLICATION NO.: PCT/AU99/00396

INTERNATIONAL FILING DATE: 19 MAY 1999

FOR: METHOD AND APPARATUS FOR TRANSFER OF REAL TIME SIGNALS OVER PACKET NETWORKS

REQUEST FOR PRIORITY UNDER 35 U.S.C. 119 AND THE INTERNATIONAL CONVENTION

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

In the matter of the above-identified application for patent, notice is hereby given that the applicant claims as priority:

COUNTRY	APPLICATION NO	DAY/MONTH/YEAR
AUSTRALIA	**PP 3624**	**19 MAY 1998**

Certified copies of the corresponding Convention application(s) were submitted to the International Bureau in PCT Application No. **PCT/AU99/00396.**

Respectfully submitted,
OBLON, SPIVAK, McCLELLAND,
MAIER & NEUSTADT, P.C.

Marvin J. Spivak
Attorney of Record
Registration No. 24,913
William E. Beaumont
Registration No. 30,996
ECKHARD H. KUESTERS
REGISTRATION NUMBER 28,870

Crystal Square Five
Fourth Floor
1755 Jefferson Davis Highway
Arlington, Virginia 22202
(703) 413-3000

gp

PCT/AU99/00396

09/462562

#4

REC'D 30 JUN 1999

WIPO PCT

Patent Office
Canberra

I, KIM MARSHALL, MANAGER EXAMINATION SUPPORT AND SALES, hereby certify that the annexed is a true copy of the Provisional specification in connection with Application No. PP 3624 for a patent by CURTIN UNIVERSITY OF TECHNOLOGY filed on 19 May 1998.

COMMONWEALTH OF AUSTRALIA
AUSTRALIA
PATENT OFFICE

WITNESS my hand this Twenty-third day of June 1999

KIM MARSHALL
MANAGER EXAMINATION SUPPORT AND SALES

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

APPLICANT: CURTIN UNIVERSITY OF TECHNOLOGY

NUMBER:

FILING DATE:

AUSTRALIA

PATENTS ACT 1990

PROVISIONAL SPECIFICATION

FOR THE INVENTION ENTITLED:

"METHOD AND APPARATUS FOR TRANSFER OF REAL TIME SIGNALS OVER PACKET NETWORK"

The invention is described in the following statement:-

Field of the Invention.

The invention is in the field of telecommunications, specifically telecommunications on a packet-switched network. It is of protocol, algorithm, procedure, and embodying network apparatus, for timely transfer and time-faithful reconstruction of real time signals, carried over the network as digital data packets. It has application in local, metropolitan and wide area networking as well as internetworking on diverse networks. It also has application in computer, control, weapon, automation, robotics and related systems.

Background of the Invention

Transfer of signals as digital data packets is already an accepted practice in computer networks. An outstanding example computer network is the world-wide Internet over which diverse signals are transferred as digital data packets, including real time signals such as telephone voice.

The transfer of real time signals over computer networks as they are at present, including the Internet, is not however satisfactory in more than one respect and for more than one reason. The transfer of data packets over a computer network is in connectionless mode, which means that packets are routed independently of each other, with the consequence that different packets carrying pieces of the same signal can reach the destination in a different sequence from the one that they left the source. Any transposition of signal segments seriously impairs the overall signal. There is no assurance that all packets sent from a source will reach the destination. Deletion signal segments impair the overall signal no less than transposition of their sequence. Delays of packets, incidence of transpositions of packets, and losses of packets depend on network loading and all become severe when the network is congested. The network does not have any arrangements or means to prevent congestion.

The invention is of protocol at the packet level, of packet assembly and dispatch procedures at source terminal, of packet routing and forwarding procedures in network elements, and of procedures for signal presentation at the destination terminal that will provide satisfactory, faithful and reliable transfer of real time signals over a packet network. The protocol of our invention carries timing information that will support the time-aware forwarding function in network routers as well as faithful presentation of signals at the receiver. It also will support synchronization at the receiver in the presentation of multiple media.

A condition for timely transfer of packets over a network is that all packets carrying a particular real time signal take the same path over the network and that elements along the path be prepared to perform the timely forwarding service on those packets. This condition amounts to a requirement that the transfer be over a particular connection and that the elements along it have reserved resources for forwarding its declared traffic. Establishing such connection-oriented transfer, complete with necessary link bandwidth and buffer resources, is in part of the known art and not of our invention. Thus, while the Internet generally has connectionless Internet Protocol datagrams, it does have a defined

resource reservation protocol (RSVP) that can fulfill the function of establishing connections with reserved bandwidth and processing resources[1].

In arriving at our invention we have recognized that, while reserved bandwidth is necessary for real time signal transfer, such reservation of bandwidth is not of itself in all cases sufficient. It is sufficient when the transfer is by small, fixed size, blocks of information at regular intervals, when the reservation of bandwidth amounts to a promise of forwarding all individual blocks of information each within a specified maximum time from its submission. But where the transfer is in large or variable length blocks of information at irregular intervals, the reservation of bandwidth can amount only to a promise of a minimum aggregated number of bits that would be transferred over an extended averaging time, such averaging time being large compared to the intervals between information blocks. The bandwidth reservation is then not sufficient to assure timely transfer. Procedures, additionally to bandwidth reservation, are then required in information forwarding elements of the network that will provide for the timely dispatch of the separate blocks of information.

It might be thought that timeliness of dispatch would be assured if the forwarding elements were sufficiently fast, work conserving, and employed first-in-first-out service disciplines. However that is adequate only when the information blocks are small and regular, and even then only when timeliness requirements are uniform across connections. Given information blocks of different sizes and/or unequal timeliness requirements among different connections through an element, the forwarding procedures need to take regard of individual requirements. It is even possible for the timeliness requirements to vary from one block to the next on the same connection, and the forwarding procedure ought to take regard of that as well.

Other disciplines besides first-in-first-out exist or are known[2], some with individual service differentiation and able to satisfy different timeliness requirements, albeit in a very limited way. A discipline that comes closest to that of our invention is the delay-earliest-due-date discipline. In this, a contract is negotiated between user and server specifying the peak and average rate of blocks, and a bound on delay that a block would need to experience in the network element. The server sets a deadline for the dispatch of a block as it enters the network element, such deadline being a fixed delay (the negotiated bound) beyond the theoretical time at which the packet would have entered, had it come in accordance with contract. There is thus possibility of individually specified timeliness, but only on a connection not individual blocks. Moreover the bandwidth reservation has to be based on the peak rate of blocks and the blocks must be of fixed size.

Insights on the fundamental requirements for real time signal transfer that have led us to our invention are that the delay between the instant at which a particular value of the signal is generated at the source and the instant at which that value is presented at the destination must be a constant over an entire communication session, and the clock by which the signal is presented at the destination must be synchronized to the clock with

1. R. Branden, L. Zhang, S. Berson, S. Herzog, "Resource Reservation Protocol (RSVP)", RFC 2205, September 1997.

2. Hui Zhang and Srinivasan Keshav "Comparison of Rate-Based Service Disciplines" Proceedings of ACM SIGCOMM'91

which the signal is generated at the source. A requirement related to the previous two and implying them, is that once presentation has commenced at the receiver, it remain continuous over the session: On the instant, or before, the last bit of the first block of information has been consumed in generating the presentation, the next block of information completes its arrival at the destination, and similarly for the second, third, and all subsequent blocks.

Further insights needed towards our invention are on the nature of representation of real time signals by digital data at varying rate, and the regeneration of the signal from those data. Blocks of data can be identified that represent particular intervals of the signal. In representation at varying rate, the size of block is not directly related to the length of interval, and vice versa. However the same length of time will elapse in the consumption of a block when generating the presentation at the receiver as took to accumulate that block at the source. Also, it is possible to represent an interval of signal by a block of data independently of the signal in other intervals or of the representation of those intervals in other blocks. We assume that such block by block independence applies to the variable rate real time signal transfer for which our invention is used.

Brief Description of the Invention

The invention is of procedures at source, at intermediate network elements, and at the destination, as well as protocol information at packet level, originating at the source, and transmitted in the packet for support of the procedures at intermediate elements and the destination.

We describe the invention in its application to the Internet, employing Internet Protocol (IP) Version 6[1]. However the invention is not restricted to the Internet or to IPv6. With appropriate modifications to protocol format, it can be applied to any packet-switched network or packet routing/switching system, whether extended over the wide area, confined to metropolitan or local areas, or even on smaller scale. Nor is it restricted to Version 6 of IP. It can be applied, with no or only obvious modification in procedures, to IPv4, the current version of the Internet Protocol.

The Internet Protocol (IP) provides the functionality for interconnecting end systems across multiple networks. To that end, IP is implemented in each end system and in routers, which correspond to the intermediate network elements of the foregoing general description. Higher level data at the source end system are encapsulated in an IP protocol data unit (PDU) for transmission. An IP PDU passes through one or more networks, and connecting routers to reach the destination end system.

For the transfer of real time signals, IP packets with appropriate real time header extension are utilized. Each IP packet in a flow represents a fragment of the real time signal. All fragments in the flow are transferred over the same Internet connection. On receiving a real time packet fragment, an IP router must forward it on time, or sooner. The

1. William Stallings "IPv6: The New Internet Protocol", IEEE Communications Magazine, July 1996, pp 96-108

time by which a fragment must be forwarded, is determined by the local state of the connection that carries the fragment. Each connecting router has to be time-aware with respect to all connections that are set up over it. Time-awareness of a router is possible only if the router has a clock, the clocks of routers and of sources and destinations are synchronous, or at least pleisochronous with each other and the router is provided with appropriate time information by sources.

The time information is carried by packet headers. We assume definition of real time extensions to IP headers. Framework for such extensions already exists. Thus in IPv6[1], it will become an instance of a hop-by-hop options header. The extension would minimally have the following three fields:

1. 2 -bit field, ORDER, indicating Beginning Of a Stream (BOS), Continuation Of a Stream (COS), End Of a Stream (EOS), Single Fragment Stream (SFS);
2. N -bit unsigned integer field, TIME, to indicate the time instant T_{abs} at the source at which the first sample of the real time signal represented by the fragment;
3. M -bit unsigned integer field, DELTA_TIME, to indicate the time interval ΔT of the real time signal that is covered by the payload of the fragment.

The size of N and M depends on the unit of time and the range of values that need to be represented. A suitable unit for time, for the purposes of our invention, is a nano-second (10^{-9} seconds).

A connection i, identified by source and destination IP addresses and UDP Port numbers or by flow label, sends fragments F_{ij}, $j = 0, 1, 2, \ldots$. The length of each fragment is denoted by L_{ij}. The timing parameter values carried by each fragment are ΔT_{ij} and $T_{abs,\,ij}$ time units.

On receipt of the first fragment F_{i0}, a router r waits for $\mathrm{WAIT}_{r,\,i}$ time units before dispatching the fragment. $\mathrm{WAIT}_{r,\,i}$ is a parameter that is fixed at connection set-up. The router r then notes the time t at which dispatch of F_{i0} was complete, calculates the deadline for the completion of dispatch of fragment F_{i1} as $\mathrm{DL}_{i1} = t + \Delta T_{i0}$ and records $T_{i,\,last} = T_{abs,\,i0} + \Delta T_{i0}$.

On receipt of the successive fragments F_{ij}, $(j > 0)$, the router r checks for a loss fragment via predicate $T_{i,\,last} \geq T_{abs,\,ij}$. If the predicate returns false indicating a loss of a fragment, the router r treats the fragment F_{ij} as F_{i0}. Otherwise, it schedules the start of dispatch as $T_{start,\,ij} = \mathrm{DL}_{ij} - L_{ij}/R_l$ where R_l is the rate of the link on which F_{ij} is to be sent. The router r dispatches F_{ij} as soon as possible, not waiting for $T_{start,\,ij}$, but observes 'earliest start - next to go' rule. The router r also calculates the deadline for the completion of dispatch of the next fragment as $\mathrm{DL}_{i(j+1)} = \mathrm{DL}_{ij} + \Delta T_{ij}$ and records $T_{i,\,last} = T_{abs,\,ij} + \Delta T_{ij}$.

On receipt of the first fragment F_{i0}, the receiver notes the time t_s and starts the presentation without delay. On finish of the presentation of F_{i0}, the receiver notes the time t_f

1. William Stallings "IPv6: The New Internet Protocol", IEEE Communications Magazine, July 1996, pp 96-108

and calculates $\Delta t = t_f - t_s$. If $\Delta t > \Delta T_{i0}$ then the application clock rate is appropriately increased. On the other hand, if $\Delta t < \Delta T_{i0}$ then the application clock rate is appropriately decreased.

On receipt of the successive fragments F_{ij}, $(j > 0)$, the receiver checks to see whether the presentation buffer is empty or not. If the buffer is empty, it restarts the presentation following an interrupt. Otherwise, the receiver waits for the presentation of previous fragments to be completed and continues the presentation without an interrupt. At the start of the presentation of F_{ij}, the receiver notes the time t_s. On finish of the presentation of F_{ij}, the receiver notes the time t_f and calculates $\Delta t = t_f - t_s$. If $\Delta t > \Delta T_{ij}$ then the application clock rate is appropriately increased. On the other hand, if $\Delta t < \Delta T_{ij}$ then the application clock rate is appropriately decreased.

With proper design and Call Admission Control (CAC), interruptions in presentation at the receiver can be expected to be rare. It can be expected that all routers, including the immediately upstream router, will have completed the dispatch of a fragment F_{ij} by their respective deadline DL_{ij}. The parameter $\mathrm{WAIT}_{r,i}$ must be chosen so that, with given CAC, it would be sufficient to ensure a low probability of missing the deadline. $\mathrm{WAIT}_{r,i}$ must not be larger than necessary as it is a component in the end-to-end delay.

Assuming no interrupts to the presentation at the receiver, fixed end-to-end delay is given by the sum of three deterministic components, namely the packetization delay ΔT_{i0} of the first fragment, sum of $\mathrm{WAIT}_{r,i}$ over all routers that the connection i traverses and the end-to-end propagation delay. For low rate applications, the packetization delay will dominate whereas, for high rate applications, the sum of $\mathrm{WAIT}_{r,i}$ will dominate.

Our invention not only provides a constant end-to-end delay of signal which is a principal consideration in real time communications but also allows the transfer of applications clock and synchronization of multiple media.